

Time-Inhomogeneous Branching Processes Conditioned on Non-Extinction

Nicholas Bhattacharya *

Mark Perlman †

Abstract

In this paper, we consider time-inhomogeneous branching processes and time-inhomogeneous birth-and-death processes, in which the offspring distribution and birth and death rates (respectively) vary in time. A classical result of branching processes states that in the critical regime, a process conditioned on non-extinction and normalized will converge in distribution to a standard exponential. In a paper of Jagers [4], time-inhomogeneous branching processes are shown to exhibit this convergence as well. In this paper, the hypotheses of Jagers' result are relaxed, further hypotheses are presented for convergence in moments, and the result is extended to the continuous-time analogue of time-inhomogeneous birth-and-death processes. In particular, the new hypotheses suggest a simple characterization of the critical regime.

Keywords: Branching processes, Birth-and-death processes, Time-inhomogeneous, Time-independent, Non-extinction

Mathematics Subject Classification (2000): 60J80

This work was started at the MAPS-REU led by Prof. D. Dolgopyat and Prof. L. Korolov at the University of Maryland in the summer of 2015.

1 Introduction

A branching process is a sequence of random variables $\{Z_n\}$ that represents a population of particles at times steps $n \in \mathbb{N}$, so that Z_n is the population of the n -th generation of particles. The initial population is one particle, $Z_0 = 1$, and subsequent generations are defined inductively. To create the generation Z_{n+1} , each of the Z_n particles reproduces independently according to a given offspring distribution, and the sum of the resulting particles quantifies the next generation Z_{n+1} . In classical, time-homogeneous branching processes, the offspring distribution X is common to all particles. In time-inhomogeneous branching processes, the offspring distribution X_n depends on the time step n . Thus a time-inhomogeneous branching process is defined by

$$\begin{cases} Z_0 = 1 \\ Z_{n+1} = \sum_{j=1}^{Z_n} X_{n,j}, \end{cases} \quad X_{n,j} \text{ are i.i.d. copies of } X_n.$$

*University of California, Berkeley. nick_bhat@berkeley.edu

†Stanford University. mpp@stanford.edu

Classical branching processes are categorized into three regimes depending on the expected number of offspring $E(X)$: a branching process is subcritical if $E(X) < 1$, critical if $E(X) = 1$ and supercritical if $E(X) > 1$. Two important results pertaining to the asymptotics of classical branching processes can be found in [1]: the first states that extinction (the event that Z_n is ever 0) occurs with probability one if and only if $E(X) \leq 1$; the second states that when a branching process in the critical regime is conditioned on non-extinction and normalized, it converges in distribution to a standard exponential:

$$\frac{\zeta_n}{E(\zeta_n)} \rightarrow \exp(1),$$

where ζ_n denotes $\{Z_n \mid Z_n \neq 0\}$, that is, Z_n conditioned on non-extinction. These results hold in the time-inhomogeneous setting, though under more complicated hypotheses. As discussed in [2], extinction occurs with probability one if and only if

$$\sum_{n=0}^{\infty} \frac{1}{E(Z_n)} = \infty,$$

under some natural regularity conditions on the offspring distribution which are thoroughly discussed in [5]. It follows from this result that the above equation naturally defines the subcritical regime for time-inhomogeneous branching processes.

The exponential limit of the conditioned, normalized process has been proven in the time-inhomogeneous setting by Jagers [4], under the main hypothesis that

$$0 < \lim_{n \rightarrow \infty} E(Z_n) < \infty,$$

along with regularity conditions. In this paper, a strictly weaker set of hypotheses is shown to be sufficient for the exponential limit, thus also providing a broader definition of the critical regime.

In the second part of this paper, the two asymptotic results mentioned above (extinction criterion and exponential limit) are reproved in the setting of time-inhomogeneous birth-and-death processes. A time-inhomogeneous birth-and-death process Z_t is a continuous-time analogue of a time-inhomogeneous branching process, in which births and deaths occur at random times determined by the current total population and in terms of infinitesimal rates:

$$\begin{cases} P(Z_{t+\Delta} - Z_t = k \mid Z_t) = Z_t \Delta b_k(t) + o(\Delta), \\ P(Z_{t+\Delta} - Z_t = 0 \mid Z_t) = 1 - Z_t \Delta (b_{-1}(t) + b_1(t) + \dots + b_n(t)) + o(\Delta). \end{cases} \quad (k = -1, 1, \dots, n)$$

The rates b_k are measurable functions defined on $[0, \infty)$. Birth-and-death processes can be considered a continuous-time analogue to branching processes, wherein individual particles reproduce or die at distinct, continuous, random times rather than reproducing all at once at every integer time. The birth rates $b_1(t), \dots, b_n(t)$ and the death rate $b_{-1}(t)$ replace the offspring distribution X_n of the discrete case. A noteworthy difference between the discrete and continuous models is that in the continuous model, the population can only increase by a fixed maximum amount at any reproduction event, whereas the support of the discrete X_n could be all of \mathbb{N} .

In [3], extinction is shown to occur with probability one under the condition

$$\int_0^{\infty} \frac{1}{E(Z_t)} dt = \infty.$$

In [6], both asymptotic results are shown in the special “fractional linear” case when there is only the single birth rate b_1 . In this paper we will extend the results to the case of finitely many birth rates.

For the remainder of the paper, the discrete setting is dealt with in its entirety before the continuous setting.

2 The Discrete Setting

2.1 Notation

We denote $g_n(s) = E(s^{X_n})$ for $s \in [0, 1]$, the generating function of the n th offspring distribution, and we denote $f_n(s) = E(s^{Z_n})$ for $s \in [0, 1]$, the generating function of the n th generation distribution. As in the case of classical branching processes, the generating function for the population can be obtained by composing generating functions of the offspring distributions:

$$f_n(s) = (g_0 \circ \dots \circ g_{n-1})(s).$$

We’ll also make use of composing only some generating functions of the offspring distributions, to obtain the generating function for $\{Z_n \mid Z_j = 1\}$, a process starting with a single particle at the j th generation:

$$f_{jn}(s) = (g_j \circ g_{j+1} \circ \dots \circ g_{n-1})(s).$$

We denote $\phi_n = P(Z_n \neq 0)$, the probability of non-extinction by the n th generation, and analogously we denote $\phi_{jn} = P(Z_n \neq 0 \mid Z_j = 1)$, the probability of non-extinction by the n th generation if we had started the process with one particle at the j th generation. We denote the means of the offspring distribution and the population distribution by

$$\begin{aligned}\mu_n &= E(X_n) = g'_n(1) \\ m_n &= E(Z_n) = f'_n(1).\end{aligned}$$

Our expression of f_n in terms of the g_n tells us that the n th population mean is the product of the preceding generation means:

$$m_n = \prod_{i=0}^{n-1} \mu_i.$$

We will more generally denote the moments and factorial moments of the population distribution by

$$\begin{aligned}M_{n,r} &= E(Z_n^r) \\ F_{n,r} &= f_n^{(r)}(1).\end{aligned}$$

We will denote by ζ_n the process Z_n conditioned on non-extinction:

$$\zeta_n = \{Z_n \mid Z_n \neq 0\}.$$

Finally, we introduce the quantity

$$\Gamma_n = \sum_{j=0}^{n-1} \frac{g_j''(1)}{2\mu_j m_{j+1}},$$

which we will show is asymptotically equal to ϕ_n^{-1} in the critical regime.

2.2 Hypotheses

We'll make use of the following hypotheses:

$$(H1) \sup_{n \geq 0} P(X_n = 0) < 1$$

$$(H2) \limsup_{n \rightarrow \infty} [P(X_n = 0) + P(X_n = 1)] < 1$$

$$(H3) \liminf_{n \rightarrow \infty} P(X_n = 0) > 0$$

$$(H4a) \sup_{n \geq 0} E(X_n^r) < \infty \text{ for } r = 1, 2, 3$$

$$(H4b) \sup_{n \geq 0} E(X_n^r) < \infty \text{ for every } r \geq 1$$

$$(H5) \Gamma_n \rightarrow \infty$$

$$(H6) m_n \Gamma_n \rightarrow \infty.$$

Hypotheses (H1)-(H4) are the regularity conditions on the offspring distribution mentioned before. Hypothesis (H1) prevents the branching process from entirely dying out in a single generation. Hypothesis (H2) prevents the possibility of every particle generating almost exactly one in the next generation, thereby causing the process to stagnate or simply stop. Hypotheses (H1) and (H2) can be replaced by the slightly stronger assumption

$$\sup_{n \geq 0} [P(X_n = 0) + P(X_n = 1)] < 1,$$

which does not allow any generations where $P(X_n = 1)$ is very close to 1, i.e. generations where almost nothing happens. Hypothesis (H3) prevents the process from tending toward a pure birth process. Hypotheses (H4a), an upper bounds on the first three moments, will be sufficient for convergence in distribution; an upper bound on all moments, (H4b), will be necessary for convergence in moments.

Hypotheses (H5) and (H6) are our new criteria for the critical regime. Hypothesis (H5) places the process out of the supercritical regime, into either the subcritical or critical regimes, and hypothesis (H6) places the process out of the subcritical regime, into either the critical or supercritical regimes. Hypothesis (H5) reduces to the familiar condition

$$\sum_{n=0}^{\infty} \frac{1}{E(Z_n)} = \infty$$

when given control on the first two moments of the offspring distribution:

$$0 < \inf_{n \geq 0} E(X_n^r) \leq \sup_{n \geq 0} E(X_n^r) < \infty \quad \text{for } r = 1, 2.$$

While it might be tempting to work with the tidier quantity

$$\sum_{j=0}^{n-1} \frac{1}{m_{j+1}} = \sum_{j=1}^n \frac{1}{E(Z_j)},$$

the forthcoming result that $\Gamma_n \phi_n \rightarrow 1$ indicates that Γ_n is the correct quantity to consider.

On necessity: the process must go extinct with probability one for the conditioned, scaled process to tend towards an exponential limit. Without this property, even a classical branching process will not exhibit the same limiting behavior in the supercritical regime. Modulo some regularity conditions, hypothesis (H5) ensures extinction. For hypothesis (H6), we rewrite the scaling factor via:

$$\begin{aligned} m_n &= E(Z_n) \\ &= E(Z_n \mid Z_n = 0)P(Z_n = 0) + E(Z_n \mid Z_n \neq 0)P(Z_n \neq 0) \\ &= E(Z_n \mid Z_n \neq 0)\phi_n. \end{aligned}$$

Thus,

$$\frac{\zeta_n}{E(Z_n \mid Z_n \neq 0)} = \frac{\zeta_n}{m_n/\phi_n}.$$

Once we show $\phi_n \Gamma_n \rightarrow 1$, hypothesis (H6) will say that the denominator in the conditioned, scaled process tends to infinity, which is certainly necessary if we are to take a discrete process ζ_n to a continuous limit $\exp(1)$. Though we do not explore it here, it is interesting to ask whether and under what hypotheses the conditioned, scaled process tends to a geometric limit in an appropriately defined subcritical regime.

On improvement: compare (H5) and (H6) with Jagers' hypothesis

$$0 < \lim_{n \rightarrow \infty} E(Z_n) < \infty.$$

Jagers' hypothesis implies (H5) and (H6), and (H5) and (H6) do not imply Jagers' hypothesis. Here we construct a simple example: let $P(X_0 = 1) = 1$, and for $n \geq 1$ let

$$P(X_n = 2) = \frac{1}{2} \frac{n+1}{n}, \quad P(X_n = 0) = 1 - P(X_n = 2).$$

Then $\mu_n = \frac{n+1}{n}$ and so $m_n = n$. This process satisfies (H1)-(H6), but does not have bounded means. Examples can be constructed similarly to attain any polynomial growth or decay of m_n , and such a process will satisfy (H5) and (H6) as long as m_n experiences polynomial decay, linear growth, or anything in between.

In addition to hypotheses (H1)-(H6), we will make use of an estimate due to Agresti [2], Lemmas 1 and 2: If $E(X_n^2) < \infty$ for each n (which is weaker than (H4a) or (H4b)), then

$$\left(\frac{1}{(1-s)m_n} + \Gamma_n \right)^{-1} \leq 1 - f_n(s) \leq \left(\frac{1}{(1-s)m_n} + \sum_{j=0}^{n-1} \frac{g_j''(0)}{2\mu_j m_{j+1}} \right)^{-1}.$$

A sharper upper bound is permitted by Agresti's argument:

$$\left(\frac{1}{(1-s)m_n} + \Gamma_n \right)^{-1} \leq 1 - f_n(s) \leq \left(\frac{1}{(1-s)m_n} + \sum_{j=0}^{n-1} \frac{g_j''(f_{jn}(0))}{2\mu_j m_{j+1}} \right)^{-1}. \quad (1)$$

Evaluating at $s = 0$ leads to

$$\left(\frac{1}{m_n} + \Gamma_n \right)^{-1} \leq \phi_n \leq \left(\frac{1}{m_n} + \sum_{j=0}^{n-1} \frac{g_j''(f_{jn}(0))}{2\mu_j m_{j+1}} \right)^{-1}. \quad (2)$$

2.3 Results

The main theorem we'll prove is the following:

Theorem 1. *If (H1)-(H6) hold, then*

$$\frac{\zeta_n}{E(\zeta_n)} \rightarrow \exp(1),$$

where the convergence is in distribution assuming (H4a), and in moments assuming (H4b).

We will prove Theorem 1 with (H4a) by showing that the moment generating function converges to that of an exponential, and we will prove Theorem 1 with (H4b) by showing that the r -th moment of the process tends to $r!$. The proofs rely on two technical lemmas and three smaller theorems.

Theorem 2. *If (H1)-(H6) hold with (H4a), then extinction occurs with probability one.*

Theorem 2 combines our hypotheses and (2) to bound $\Gamma_n \phi_n$ from above and below, showing that $\phi_n \rightarrow 0$. The proof of Theorem 2 relies almost entirely on the two preceding lemmas. Theorem 3, however, will require more careful estimates to show the asymptotic equality:

Theorem 3. *If (H1)-(H6) hold with (H4a), then $\phi_n \Gamma_n \rightarrow 1$.*

Theorems 2 and 3 are sufficient to prove Theorem 1 with (H4a). For the moments convergence with (H4b), Theorem 4 relates Γ_n to the factorial moments of Z_n .

Theorem 4. *If (H1)-(H6) hold with (H4b), then*

$$\frac{F_{n,r}}{m_n^r \Gamma_n^{r-1}} \rightarrow r! \quad (r \geq 1).$$

These are all the necessary parts to complete Theorem 1. As an endnote, we prove that hypothesis (H6) can be substituted for a stronger but more intuitive condition:

Theorem 5. *Given (H1)-(H5), if $\mu_n \rightarrow 1$ then (H6) holds.*

The difference between (H6) and the condition $\mu_n \rightarrow 1$ is that (H6) allows for generation means that oscillate around 1.

2.4 Two Technical Lemmas

First we point out a few immediate technical consequences of (H1)-(H6). Hypotheses (H1) and (H4) imply that μ_n is uniformly bounded away from zero and from infinity. Hypothesis (H2) implies that $\liminf_{n \rightarrow \infty} g_n''(1) > 0$, because $\sum_{k=2}^{\infty} P(X_n = k)$ is bounded away from zero. And finally, because factorial moments can be expressed as combinations of regular moments, we have $\sup_{n \geq 0} g_n^{(r)}(1) < \infty$ for $r = 1, 2, 3$ under (H4a) or all $r \geq 1$ under (H4b).

Lemma 1. *Hypotheses (H2) and (H4) imply that there exists an N such that for every $0 < \varepsilon_0 \leq 1$, we have*

$$\inf_{\varepsilon \geq \varepsilon_0} \inf_{n \geq N} g_n''(\varepsilon) > 0.$$

Lemma 1 says that the generating functions of the offspring distributions are convex, uniformly in n . Convexity should follow intuitively from hypothesis (H2), because

$$g_n''(\varepsilon) = \sum_{k=2}^{\infty} k(k-1)P(X_n = k)\varepsilon^{k-2},$$

and (H2) tells us that the sum of the coefficients $\{P(X_n = k)\}_{k=2}^{\infty}$ stays away from zero. To get uniformity in n , The only possible problem is mass tending towards the tails of the X_n distribution, and this is precluded by hypothesis (H4).

Proof. Since $g_n''(s)$ is a power series of nonnegative terms, it is increasing in s . Therefore the uniformity in $\varepsilon \geq \varepsilon_0$ follows immediately from showing that $\inf_{n \geq N} g_n''(\varepsilon) > 0$ for some fixed $\varepsilon > 0$.

By (H4) pick M such that $\mu_n < M$ for every n . By (H2) pick $N, \delta > 0$ such that for all $n \geq N$, we have $\sum_{k=2}^{\infty} P(X_n = k) > \delta$. Then $\sum_{k=2}^{2M/\delta} P(X_n = k) \geq \frac{\delta}{2}$ for all $n \geq N$; to see this, suppose otherwise. Then there would be an $n \geq N$ such that

$$\mu_n \geq \sum_{k=2M/\delta}^{\infty} kP(X_n = k) \geq \frac{2M}{\delta} \sum_{k=2M/\delta}^{\infty} P(X_n = k) > M,$$

a contradiction. Now we have

$$g_n''(\varepsilon) = \sum_{k=2}^{\infty} k(k-1)P(X_n = k)\varepsilon^{k-2} \geq \sum_{k=2}^{2M/\delta} P(X_n = k)\varepsilon^{2M/\delta} \geq \frac{\delta}{2} \cdot \varepsilon^{2M/\delta}.$$

□

Lemma 2. *Assumptions (H1)-(H5) imply that there exist $C, N > 0$ such that for every $n \geq N$, we have*

$$\Gamma_n \equiv \sum_{j=0}^{n-1} \frac{g_j''(1)}{2\mu_j m_{j+1}} \leq C \sum_{j=0}^{n-1} \frac{g_j''(f_{jn}(0))}{2\mu_j m_{j+1}}.$$

Lemma 2 is a first, rough relation between the upper bound in (2) and Γ_n . Note that $f_{jn}(0)$ is simply $1 - \phi_{jn}$, the probability of the process started at time step j dying out by time n . Because extinction is sure, we know that $\phi_{jn} \rightarrow 0$ as $n - j \rightarrow \infty$, and so “most” of the terms of the sum on the right will have $f_{jn}(0)$ close to 1.

Proof. Hypothesis (H5) tells us that all of the mass of Γ_n is in the tail of the sum, and so it suffices to show that there exist $N, C > 0$ such that $g_j''(1) \leq C g_j''(f_{jn}(0))$ for all $n \geq N$ and $j < n$. By (H3), fix N such that $\inf_{n \geq N} g_n(0) > 0$. Now for any $j \geq N$, f_{jn} is a composition of g_k ’s satisfying $\inf_{k \geq N} \|g_k\|_\infty > 0$, because every g_k is an increasing function and bounded uniformly below at 0. Thus we have

$$\inf_{n > N} \inf_{N \leq j < n} f_{jn}(0) > 0.$$

If, on the other hand, $j < N$, then we write

$$f_{jn}(0) = (f_{jN} \circ f_{Nn})(0).$$

From the previous case, we know that $f_{Nn}(0)$ is bounded away from 0 uniformly. From (H1), we know that the g_j are all strictly increasing and nonnegative, and so f_{jN} is a composition of at most N functions that are positive away from 0. Thus once again,

$$\inf_{n > N} \inf_{j < n} f_{jn}(0) > 0.$$

Thus for $n \geq N$, the argument of g_j'' is bounded away from 0, and so by Lemma 1, picking N larger if necessary, we have $\inf_{n \geq N} g_j''(f_{jn}(0)) > 0$. From (H4) we know that $\sup_{n \geq 0} g_j''(1) < \infty$, and so we can bound the latter quantity uniformly by the former. Since $\Gamma_n \rightarrow \infty$, we can ignore the first N terms of the sum and substitute this bound in the sum, proving the desired result. \square

2.5 Proof of Theorem 2

To prove Theorem 2, we simply put together the lemmas we’ve shown, hypothesis (H6), and the estimate (2).

Proof. Consider (2). Hypothesis (H6) tells us that the left hand side is asymptotically $\frac{1}{\Gamma_n}$. Hypothesis (H6) and Lemma 2 tell us that the right hand side is asymptotically smaller than $\frac{C}{\Gamma_n}$ for some constant C . Consequently,

$$1 \leq \liminf_{n \rightarrow \infty} \phi_n \Gamma_n \leq \limsup_{n \rightarrow \infty} \phi_n \Gamma_n \leq C.$$

Since $\Gamma_n \rightarrow \infty$, we have $\phi_n \rightarrow 0$ and thus extinction occurs with probability one. \square

2.6 Proof of Theorem 3

The proof of Theorem 3 can be thought of as a more careful version of Lemma 2. In the sum involving $g_n''(f_{jn}(0))$, we show that the terms in which $f_{jn}(0)$ is far away from 1 contribute a negligible amount to the sum, so that the constant C in Lemma 2 is in fact equal to 1.

Proof. 1. Applying the inequality $\phi_n \leq \frac{C}{\Gamma_n}$ to the process started at time j (that is, $\{Z_n|Z_j = 1\}$) gives us

$$\phi_{jn} \leq C \frac{1}{\sum_{k=j}^{n-1} \frac{g_k''(1)}{2\mu_k(m_{k+1}/m_j)}} = C \frac{1}{m_j(\Gamma_n - \Gamma_j)}.$$

Note that m_k/m_j is the mean of the process starting at time j .

2. By (H3), pick N large enough and $\eta > 0$ small enough that $\inf_{n \geq N} g_n''(1) > \eta$. By (H4), pick L such that $\sup_{n \geq 0} g_n'''(1) < L$. Let $\varepsilon > 0$ and define the index

$$J(n) = \min\{j : \phi_{jn} > \varepsilon\eta/L\}.$$

The indices satisfying $j < J(n)$ correspond to the terms of ϕ_{jn} that are very close to 0 (i.e. terms where $f_{jn}(0)$ is very close to 1). We will show that “most” terms satisfy $j < J(n)$, in the sense that we will show that the terms with $j \geq J(n)$ contribute negligibly to the summation in the upper bound of (2).

Since extinction is sure, $\phi_{jn} \rightarrow 0$ for fixed j as $n \rightarrow \infty$. Consequently, $J(n) \rightarrow \infty$ as $n \rightarrow \infty$. If $N \leq j < J(n)$, then $\phi_{jn} \leq \varepsilon\eta/L$, and we calculate via the mean value theorem:

$$\begin{aligned} g_j''(f_{jn}(0)) &= g_j''(1 - \phi_{jn}) \\ &= g_j''(1) - g_j'''(\xi)\phi_{jn}, \quad \xi \in (1 - \phi_{jn}, 1) \\ &\geq g_j''(1) - L\phi_{jn} \\ &= g_j''(1) \left(1 - \frac{L\phi_{jn}}{g_j''(1)}\right) \\ &\geq g_j''(1) \left(1 - \frac{L\phi_{jn}}{\eta}\right) \\ &\geq g_j''(1)(1 - \varepsilon). \end{aligned}$$

3. Using (2) and our result in 2., we find at large n :

$$\begin{aligned} 1 &\geq \frac{1}{\phi_n \Gamma_n} \geq \frac{1}{\Gamma_n} \left(\frac{1}{m_n} + \sum_{j=1}^{n-1} \frac{g_j''(f_{jn}(0))}{2\mu_j m_{j+1}} \right) \\ &\geq \frac{1}{\Gamma_n} \sum_{j=1}^{J(n)-1} \frac{g_j''(f_{jn}(0))}{2\mu_j m_{j+1}} \\ &\geq \frac{1}{\Gamma_n} \sum_{j=1}^{J(n)-1} \frac{g_j''(1)(1 - \varepsilon)}{2\mu_j m_{j+1}} \\ &\geq \frac{1}{\Gamma_n} (1 - \varepsilon) \Gamma_{J(n)}. \end{aligned}$$

Although our result in 2. also requires $N \leq j$, we can ignore finitely many of the terms of the series because of hypothesis (H5). Next, using the fact that $\Gamma_{J(n)} \leq \Gamma_n$,

$$\begin{aligned} \frac{1}{\Gamma_n}(1 - \varepsilon)\Gamma_{J(n)} &= \frac{\Gamma_{J(n)}}{\Gamma_n} - \varepsilon \frac{\Gamma_{J(n)}}{\Gamma_n} \\ &\geq \frac{\Gamma_{J(n)}}{\Gamma_n} - \varepsilon \\ &= (1 - \varepsilon) - \frac{\Gamma_n - \Gamma_{J(n)}}{\Gamma_n}. \end{aligned}$$

We want to show that the rightmost term goes to zero. Our result in 1. and the definition of $J(n)$ give us

$$\frac{\Gamma_n - \Gamma_{J(n)}}{\Gamma_n} \leq \frac{C}{m_{J(n)}\phi_{J(n)n}\Gamma_n} \leq \frac{CL}{\varepsilon\eta} \cdot \frac{1}{m_{J(n)}\Gamma_n}.$$

Since $J(n) \rightarrow \infty$, we have

$$m_{J(n)}\Gamma_n \geq m_{J(n)}\Gamma_{J(n)} \rightarrow \infty,$$

and thus the desired term goes to zero. Whence,

$$1 \geq \limsup_{n \rightarrow \infty} \frac{1}{\phi_n\Gamma_n} \geq \liminf_{n \rightarrow \infty} \frac{1}{\phi_n\Gamma_n} \geq 1 - \varepsilon,$$

for every $\varepsilon > 0$, completing the proof. □

2.7 Proof of Theorem 4

Now we move our attention to the factorial moments of Z_n , and their relation to Γ_n . In addition to the presence of Γ_n in (2), the proof of Theorem 4 offers some insight into how Γ_n arises as an important quantity. At this point, we start assuming (H4b).

Proof. 1. Taking the r -th derivative of $f_{n+1}(s) = f_n(g_n(s))$ and evaluating at 1, we find

$$F_{n+1,r} = \mu_n^r F_{n,r} + C_{n,r-1}F_{n,r-1} + \dots + C_{n,1}F_{n,1},$$

where $C_{n,j}$ is a coefficient composed of factorial moments of X_n and by (H4b) is thus bounded uniformly in n for each j . The only coefficient we'll need explicitly is $C_{n,r-1} = \binom{r}{2}g_n''(1)\mu_n^{r-2}$. Dividing the above quantity by m_{n+1} , we find a recursion relation for $F_{n,r}$ on n , which we then iterate:

$$\begin{aligned} \frac{F_{n+1,r}}{m_{n+1}^r} &= \frac{F_{n,r}}{m_n^r} + C_{n,r-1} \frac{F_{n,r-1}}{m_{n+1}^r} + \dots + C_{n,1} \frac{F_{n,1}}{m_{n+1}^r} \\ &= \frac{F_{n,r}}{m_n^r} + \sum_{k=1}^{r-1} \frac{C_{n,k}F_{n,k}}{m_{n+1}^r} \\ &= \sum_{k=1}^{r-1} \sum_{j=0}^n \frac{C_{j,k}F_{j,k}}{m_{j+1}^r}. \end{aligned}$$

Then, expanding the sum over the other index and splitting off some of the factors of m_{j+1} in the denominator,

$$\begin{aligned}
\frac{F_{n+1,r}}{m_{n+1}^r} &= \sum_{j=0}^n \sum_{k=1}^{r-1} \frac{C_{j,k} F_{j,k}}{m_{j+1}^r} \\
&= \sum_{j=0}^n C_{j,r-1} \frac{F_{j,r-1}}{m_{j+1}^r} + \dots + \sum_{j=0}^n C_{j,1} \frac{F_{j,1}}{m_{j+1}^r} \\
&= \sum_{j=0}^n \frac{C_{j,r-1}}{m_j \mu_j^r} \frac{F_{j,r-1}}{m_j^{r-1}} + \dots + \sum_{j=0}^n \frac{C_{j,1}}{m_j^{r-1} \mu_j^r} \frac{F_{j,1}}{m_j}. \tag{3}
\end{aligned}$$

2. We will prove Theorem 4 inductively. We know that $F_{n,1} = m_n$, so the claim is clear for $r = 1$. Assume inductively it's true for $k = 1, \dots, r-1$, so that

$$\frac{F_{n,k}/m_n^k}{k! \Gamma_n^{k-1}} \xrightarrow{n \rightarrow \infty} 1, \quad (k = 1, \dots, r-1),$$

and also

$$\frac{\frac{C_{n,k}}{m_n^{r-k} \mu_n^r} F_{n,k}/m_n^k}{\frac{C_{n,k}}{m_n^{r-k} \mu_n^r} k! \Gamma_n^{k-1}} \xrightarrow{n \rightarrow \infty} 1, \quad (k = 1, \dots, r-1).$$

We aim to substitute this into (3). Recall the general fact that if $\frac{a_n}{b_n} \rightarrow 1$ and b_n is not summable, then $\frac{\sum_{j=1}^n a_j}{\sum_{j=1}^n b_j} \rightarrow 1$ as well. We would like to use this fact to say

$$\frac{\sum_{j=0}^n \frac{C_{j,k}}{m_j^{r-k} \mu_j^r} F_{j,k}/m_j^k}{\sum_{j=0}^n \frac{C_{j,k}}{m_j^{r-k} \mu_j^r} k! \Gamma_j^{k-1}} \rightarrow 1, \quad (k = 1, \dots, r-1),$$

but we can only use this relation in the largest few values of k , because $\frac{C_{n,k}}{m_n^{r-k} \mu_n^r} k! \Gamma_n^{k-1}$ is not summable for $k = r-1$ (corresponding to the leading series in (3)) and it is summable for $k = 1$ (corresponding to the last series in (3)). But since we're adding all these sums together, we only care about the ones that tend to infinity, so we'll clump the finite sums into an $O(1)$, and use the fact for the rest:

$$\begin{aligned}
\frac{F_{n+1,r}}{m_{n+1}^r} &= \left[\sum_{j=0}^n \frac{C_{j,r-1}}{m_j \mu_j^r} (r-1)! \Gamma_j^{r-2} + o \left(\sum_{j=0}^n \frac{C_{j,r-1}}{m_j \mu_j^r} (r-1)! \Gamma_j^{r-2} \right) \right] \\
&\quad + \left[\sum_{j=0}^n \frac{C_{j,r-2}}{m_j \mu_j^r} (r-2)! \Gamma_j^{r-3} + o \left(\sum_{j=0}^n \frac{C_{j,r-2}}{m_j \mu_j^r} (r-2)! \Gamma_j^{r-3} \right) \right] + \dots + O(1) \\
&= \left[\sum_{j=0}^n \frac{(r-1)! C_{j,r-1}}{\mu_j^r} \frac{\Gamma_j^{r-2}}{m_j} + o \left(\sum_{j=0}^n \frac{\Gamma_j^{r-2}}{m_j} \right) \right] \\
&\quad + \left[\sum_{j=0}^n \frac{(r-2)! C_{j,r-2}}{\mu_j^r} \frac{\Gamma_j^{r-3}}{m_j^2} + o \left(\sum_{j=0}^n \frac{\Gamma_j^{r-3}}{m_j^2} \right) \right] + \dots + O(1).
\end{aligned}$$

3. The growth of the above sums is driven by the $\frac{\Gamma_j}{m_j}$ term. Since $m_n \Gamma_n \rightarrow \infty$, the first sum is of higher order than the others, and thus

$$\begin{aligned} \frac{F_{n+1,r}}{m_{n+1}^r} &= \sum_{j=0}^n \frac{(r-1)! C_{j,r-1}}{\mu_j^r} \frac{\Gamma_j^{r-2}}{m_j} + o\left(\sum_{j=0}^n \frac{\Gamma_j^{r-2}}{m_j}\right) \\ &= r! \sum_{j=0}^n \frac{(r-1)g_j''(1)}{2\mu_j} \frac{\Gamma_j^{r-2}}{m_{j+1}} + o\left(\sum_{j=0}^n \frac{\Gamma_j^{r-2}}{m_j}\right), \end{aligned}$$

where in the last line we've substituted for $C_{j,r-1}$. To complete the inductive proof we just need to show

$$\Gamma_{n+1}^{r-1} = \sum_{j=0}^n \frac{(r-1)g_j''(1)}{2\mu_j} \frac{\Gamma_j^{r-2}}{m_{j+1}} + o\left(\sum_{j=0}^n \frac{\Gamma_j^{r-2}}{m_{j+1}}\right).$$

To simplify notation, we show it for $r \mapsto r+1$; that is, we want to show

$$\Gamma_{n+1}^r = \sum_{j=0}^n \frac{r g_j''(1)}{2\mu_j} \frac{\Gamma_j^{r-1}}{m_{j+1}} + o\left(\sum_{j=0}^n \frac{\Gamma_j^{r-1}}{m_{j+1}}\right).$$

4. A binomial expansion of Γ_{n+1}^r gives us another recurrence relation:

$$\begin{aligned} \Gamma_{n+1}^r &= \left(\Gamma_n + \frac{g_n''(1)}{2\mu_n m_{n+1}}\right)^r \\ &= \sum_{j=0}^r \binom{r}{j} \left(\frac{g_n''(1)}{2\mu_n m_{n+1}}\right)^j \Gamma_n^{r-j} \\ &= \Gamma_n^r + r \frac{g_n''(1)}{2\mu_n} \frac{\Gamma_n^{r-1}}{m_{n+1}} + \dots + \left(\frac{g_n''(1)}{2\mu_n}\right)^r \frac{1}{m_{n+1}^r}. \end{aligned}$$

Expanding that recurrence relation yields

$$\Gamma_{n+1}^r = \sum_{j=0}^n \frac{r g_j''(1)}{2\mu_j} \frac{\Gamma_j^{r-1}}{m_{j+1}} + \dots + \sum_{j=0}^n \left(\frac{g_j''(1)}{2\mu_j}\right)^r \frac{1}{m_{j+1}^r}.$$

As before, the first sum dominates the others, and we have

$$\Gamma_{n+1}^r = \sum_{j=0}^n \frac{r g_j''(1)}{2\mu_j} \frac{\Gamma_j^{r-1}}{m_{j+1}} + o\left(\sum_{j=0}^n \frac{\Gamma_j^{r-1}}{m_{j+1}}\right),$$

as desired. □

2.8 Proof of Theorem 1

Proving Theorem 1 is now just a matter of putting together the pieces of the other theorems.

Proof. 1. Assume (H4a). We aim to show that the moment generating function of $\frac{\zeta_n}{E(\zeta_n)}$ converges to $\frac{1}{1-s}$. First we expand the generating function of Z_n by:

$$\begin{aligned} f_n(s) &= E(s^{Z_n}) = P(Z_n \neq 0)E(s^{Z_n}|Z_n \neq 0) + P(Z_n = 0)E(s^{Z_n}|Z_n = 0) \\ &= \phi_n E(s^{Z_n}|Z_n \neq 0) + (1 - \phi_n). \end{aligned}$$

Then we substitute $s = \exp(-\frac{s}{E(Z_n|Z_n \neq 0)})$ to obtain the moment generating function of $-s$, which we will show tends to $\frac{1}{1+s}$. Also substituting $E(Z_n|Z_n \neq 0) = m_n/\phi_n$,

$$\begin{aligned} E\left(\exp\left(-\frac{sZ_n}{E(Z_n|Z_n \neq 0)}\right) \middle| Z_n \neq 0\right) &= \frac{f_n(\exp(-\frac{s}{E(Z_n|Z_n \neq 0)})) - 1}{\phi_n} + 1 \\ &= \frac{f_n(\exp(-\frac{s\phi_n}{m_n})) - 1}{\phi_n} + 1. \end{aligned}$$

2. Part 3 of the proof of Theorem 3 showed that

$$\sum_{j=0}^{n-1} \frac{g_j''(f_{jn}(0))}{2\mu_j m_{j+1}} = \Gamma_n + o(\Gamma_n),$$

and thus (1) takes the form

$$\left(\frac{1}{(1-s)m_n} + \Gamma_n\right)^{-1} \leq 1 - f_n(s) \leq \left(\frac{1}{(1-s)m_n} + \Gamma_n + o(\Gamma_n)\right)^{-1}.$$

Cleaning up this expression to make it look more like our moment generating function, we find

$$1 - \frac{1}{\phi_n} \left(\frac{1}{(1-s)m_n} + \Gamma_n + o(\Gamma_n)\right)^{-1} \leq \frac{f_n(s) - 1}{\phi_n} + 1 \leq 1 - \frac{1}{\phi_n} \left(\frac{1}{(1-s)m_n} + \Gamma_n\right)^{-1}.$$

And substituting $s = \exp(-\frac{s\phi_n}{m_n})$,

$$\begin{aligned} 1 - \frac{1}{\phi_n} \left(\frac{1}{(1 - \exp(-\frac{s\phi_n}{m_n}))m_n} + \Gamma_n + o(\Gamma_n)\right)^{-1} &\leq E\left(\exp\left(-\frac{sZ_n}{E(Z_n|Z_n \neq 0)}\right) \middle| Z_n \neq 0\right) \\ &\leq 1 - \frac{1}{\phi_n} \left(\frac{1}{(1 - \exp(-\frac{s\phi_n}{m_n}))m_n} + \Gamma_n\right)^{-1}. \end{aligned}$$

Thus it suffices to show that

$$1 - \frac{1}{\phi_n} \left(\frac{1}{(1 - e^{-\frac{s\phi_n}{m_n}})m_n} + \Gamma_n\right)^{-1} \rightarrow \frac{1}{1+s}.$$

3. Hypothesis (H6) and Theorem 3 imply that $\frac{\phi_n}{m_n} \rightarrow 0$, and so $e^{-\frac{s\phi_n}{m_n}} \rightarrow 1$. Taking a first order approximation and using Theorem 3,

$$\begin{aligned}
1 - \frac{1}{\phi_n} \left(\frac{1}{(1 - e^{-\frac{s\phi_n}{m_n}})m_n} + \Gamma_n \right)^{-1} &= 1 - \frac{1}{\phi_n} \left(\frac{1}{(s\phi_n/m_n + o(\phi_n/m_n))m_n} + \Gamma_n \right)^{-1} \\
&= 1 - \frac{1}{\Gamma_n \phi_n} \left(\frac{1}{s\phi_n \Gamma_n + o(\phi_n \Gamma_n)} + 1 \right)^{-1} \\
&\rightarrow 1 - \frac{s}{1+s} \\
&= \frac{1}{1+s},
\end{aligned}$$

as desired. This concludes the proof in the case of (H4a).

4. Now assume (H4b). The r -th moment of the conditioned, scaled process can be written as

$$\begin{aligned}
E \left(\left(\frac{Z_n}{E(Z_n|Z_n \neq 0)} \right)^r \middle| Z_n \neq 0 \right) &= \frac{E(Z_n^r|Z_n \neq 0)}{E(Z_n|Z_n \neq 0)^r} \\
&= \frac{M_{n,r}/\phi_n}{(m_n/\phi_n)^r} \\
&= \frac{M_{n,r}\phi_n^{r-1}}{m_n^r}.
\end{aligned}$$

Theorem 3 tells us

$$\lim_{n \rightarrow \infty} \frac{M_{n,r}\phi_n^{r-1}}{m_n^r} = \lim_{n \rightarrow \infty} \frac{M_{n,r}}{m_n^r \Gamma_n^{r-1}}.$$

We can write moments in terms of factorial moments by

$$M_{n,r} = \sum_{k=0}^r \left\{ \begin{matrix} r \\ k \end{matrix} \right\} F_{n,k},$$

where $\left\{ \begin{matrix} r \\ k \end{matrix} \right\}$ is a Stirling number of the second kind. Theorem 4 and hypothesis (H5) tell us that the highest-index term dominates the sum, and since $\left\{ \begin{matrix} r \\ r \end{matrix} \right\} = 1$, we find

$$\lim_{n \rightarrow \infty} \frac{M_{n,r}}{m_n^r \Gamma_n^{r-1}} = \lim_{n \rightarrow \infty} \frac{F_{n,r}}{m_n^r \Gamma_n^{r-1}} = r!$$

as desired, completing the proof. □

2.9 Proof of Theorem 5

Proof. Hypotheses (H1)-(H4) let us write

$$\Gamma_n C \geq \sum_{j=0}^{n-1} \frac{1}{m_{j+1}} \geq C' \Gamma_n.$$

for constants C, C' . Since $m_{j+1} = \mu_j m_j$, we can write

$$\frac{1}{m_n} = 1 + \sum_{j=0}^{n-1} \left(\frac{1}{m_{j+1}} - \frac{1}{m_j} \right) = 1 + \sum_{j=0}^n \frac{1 - \mu_j}{m_{j+1}}.$$

Fix $\varepsilon > 0$ and pick N large enough that $|1 - \mu_n| < \varepsilon$ for all $n > N$. Then when $n > N$, combining these inequalities yields

$$\begin{aligned} \frac{1}{m_n \Gamma_n} &\leq C \frac{1 + \sum_{j=0}^{n-1} \frac{|1 - \mu_j|}{m_{j+1}}}{\sum_{j=0}^{n-1} \frac{1}{m_{j+1}}} \\ &= \frac{C}{\sum_{j=0}^{n-1} \frac{1}{m_{j+1}}} + \frac{C \sum_{j=0}^N \frac{|1 - \mu_j|}{m_{j+1}}}{\sum_{j=0}^{n-1} \frac{1}{m_{j+1}}} + \frac{C \sum_{j=N+1}^{n-1} \frac{|1 - \mu_j|}{m_{j+1}}}{\sum_{j=0}^{n-1} \frac{1}{m_{j+1}}} \\ &\leq \frac{C}{\sum_{j=0}^{n-1} \frac{1}{m_{j+1}}} + \frac{C \sum_{j=0}^N \frac{|1 - \mu_j|}{m_{j+1}}}{\sum_{j=0}^{n-1} \frac{1}{m_{j+1}}} + C\varepsilon. \end{aligned}$$

Hypothesis (H5) tells us that the first two sums tend to zero, and thus the entire last expression tends to $C\varepsilon$. Thus $\frac{1}{m_n \Gamma_n} \rightarrow 0$, as desired. \square

This concludes the discrete setting of branching processes. Now we move on to the continuous setting of birth-and-death processes.

3 The Continuous Setting

The continuous setting will be presented and proved in exactly the same manner that the discrete setting was. However, the discrete setting relied heavily on generating function relations that are not present in the continuous setting, so the continuous setting will require different techniques.

3.1 Notation

We will often drop the time argument of rates and means when it is implied.

We define the function X_t by

$$\begin{aligned} -1 &\mapsto b_{-1}(t) \\ 1 &\mapsto b_1(t) \\ &\vdots \\ n &\mapsto b_n(t). \end{aligned}$$

In this sense, X_t is similar to a random variable that has probability $b_k(t)$ of taking value k ; however, $\sum_{k=-1,1,\dots,n} b_k(t)$ need not equal 1, so X_t is not an actual random variable. For purely notational

reasons, though, it will be useful to denote moments of X_t as though it were a random variable; thus we denote

$$\begin{aligned} E(X_t) &= \sum_{k=-1,1,\dots,n} k b_k(t) \\ E(X_t^r) &= \sum_{k=-1,1,\dots,n} k^r b_k(t). \end{aligned}$$

We define $b_0 = -\sum_{k=-1,1,\dots,n} b_k$, which gives us a cleaner expression for transition probabilities:

$$\begin{cases} P(Z_{t+\Delta} - Z_t = k \mid Z_t) = Z_t \Delta b_k(t) + o(\Delta), & (k = -1, 1, \dots, n) \\ P(Z_{t+\Delta} - Z_t = 0 \mid Z_t) = 1 + Z_t \Delta b_0(t) + o(\Delta). \end{cases}$$

In the same way that X_t is an analogue of X_n in the discrete case, we introduce the function

$$g_t(x) = \sum_{k=-1}^n b_k(t) x^{k+1},$$

which is (similarly to X_t) not a generating function, because the coefficient b_0 is negative. It will nonetheless prove a useful notation for our needs. Indeed, we remark immediately that

$$\begin{aligned} g_t(1) &= 0 \\ g'_t(1) &= E(X_t) \\ g''_t(1) &= E(X_t^2) + E(X_t) \\ g'''_t(1) &= E(X_t^3) - E(X_t). \end{aligned}$$

We denote the moments of Z_t by

$$\begin{aligned} M(t) &= E(Z_t) \\ M_r(t) &= E(Z_t^r) \end{aligned}$$

and probabilities of survival by

$$\begin{aligned} \phi(t) &= P(Z_t \neq 0) \\ \phi(s, t) &= P(Z_t \neq 0 \mid Z_s = 1). \end{aligned}$$

Once again, we will denote by ζ_t the process conditioned on non-extinction:

$$\zeta_t = \{Z_t \mid Z_t \neq 0\}.$$

Finally, we introduce our continuous analogue of Γ :

$$\Gamma(t) = \int_0^t \frac{E(X_s^2)}{2M(s)} ds,$$

which will once again be shown to satisfy $\Gamma(t)\phi(t) \rightarrow 1$ in the critical regime.

3.2 Hypotheses

We'll make use of the following hypotheses:

$$(H7) \quad \liminf_{t \geq 0} b_{-1}(t) > 0$$

$$(H8) \quad \liminf_{t \geq 0} b_1(t) + \dots + b_n(t) > 0$$

$$(H9) \quad \sup_{t \geq 0} [b_{-1}(t) + b_1(t) + \dots + b_n(t)] < \infty$$

$$(H10) \quad \Gamma(t) \rightarrow \infty$$

$$(H11) \quad M(t)\Gamma(t) \rightarrow \infty.$$

Hypotheses (H7)-(H9) are the analogous regularity conditions. Hypotheses (H7) and (H8) prevent the process from tending towards a pure birth process or pure death process, respectively.

Hypothesis (H9) bounds all moments of X_t , uniformly in t . Here we see an implication of allowing only finitely many births: there is no option to bound just finitely many moments and prove that the conditioned, scaled process converges in distribution. Indeed, our analogue to Theorem 1 will only deal with convergence in moments.

Hypothesis (H9) also assures us the ability to approximate the b_k by continuous functions with arbitrarily small L^1 error. Lemma 1 will show that we only need to care about the L^1 behavior of the rates and not the pointwise behavior. Consequently, we will often treat the b_k as though they were continuous to differentiate integrals involving the b_k .

Hypotheses (H10) and (H11) are exactly analogous to the discrete setting: hypothesis (H10) places the process out of the supercritical regime, and hypothesis (H11) places the process out of the subcritical regime. Under our regularity assumptions, hypothesis (H10) is actually equivalent to the familiar condition

$$\int_0^\infty \frac{1}{E(Z_t)} dt = \infty.$$

Hypotheses (H10) and (H11) are necessary for exactly the same reasons that they are necessary in the discrete setting.

In addition to hypotheses (H7)-(H11), we will make use of an estimate due to Cistjakov and Markova [3], Theorem 1:

$$\phi(t) = \frac{1}{\frac{1}{M(t)} + \int_0^t \frac{g_s(1 - \phi(s, t)) + \phi(s, t)g'_s(1)}{\phi(s, t)^2 M(s)} ds}.$$

A second order expansion of g_s (a sharper version of [3], Theorem 2) yields

$$\frac{1}{\frac{1}{M(t)} + \int_0^t \frac{g''_s(1)}{2M(s)} ds} \leq \phi(t) \leq \frac{1}{\frac{1}{M(t)} + \int_0^t \frac{g''_s(1 - \phi(s, t))}{2M(s)} ds}. \quad (4)$$

3.3 Results

The main theorem we'll prove is the following:

Theorem 6. *If (H7)-(H11) hold, then*

$$\frac{\zeta_t}{E(\zeta_t)} \rightarrow \exp(1),$$

where the convergence is in moments.

The proof relies on four lemmas and three smaller theorems.

Theorem 7. *If (H7)-(H11) hold, then extinction occurs with probability one.*

Theorem 7 will be proven in the same way that Theorem 2 was in the discrete setting, and likewise, Theorem 8 will be proven as Theorem 3 was.

Theorem 8. *If (H7)-(H11) hold, then $\phi(t)\Gamma(t) \rightarrow 1$.*

Whereas Theorem 4 was proven using recurrence relations between generating functions, Theorem 9 will be proven using recurrence relations between the moments of Z_t ; specifically, the relations in the forthcoming Lemma 3.

Theorem 9. *If (H7)-(H11) hold, then*

$$\frac{M_r(t)}{M(t)^r \Gamma(t)^{r-1}} \rightarrow r! \quad (r \geq 1).$$

Theorems 7-9 will be enough to prove Theorem 6. Once again, we end with an additional theorem that allows us to substitute hypothesis (H11) for a stronger but more intuitive condition:

Theorem 10. *Given (H7)-(H10), if $E(X_t) \rightarrow 0$ then (H11) holds.*

3.4 Four Lemmas

Lemma 3.

$$\dot{M}_r(t) = \sum_{j=1}^r \binom{r}{j} M_{r-j+1}(t) E(X_t^j).$$

This recurrence relation between the moments of the process will replace the recurrence relation between generating functions in the discrete setting. Lemma 3 implies in particular that $M(t) = \exp(\int_0^t E(X_s) ds)$, an equality we used when transcribing (4).

Proof. The probability of two changes in Z_t happening in a time span of Δ is $o(\Delta)$, so we calculate

$$\begin{aligned} E(Z_{t+\Delta}^r | Z_t) &= \sum_{k=-1}^n (Z_t + k)^r P(Z_{t+\Delta} = Z_t + k | Z_t) + o(\Delta) \\ &= Z_t^r (1 + Z_t \Delta b_0) + \sum_{k=-1, 1, \dots, n} (Z_t + k)^r Z_t \Delta b_k + o(\Delta) \\ &= Z_t + Z_t \Delta \sum_{k=-1}^n (Z_t + k)^r b_k + o(\Delta). \end{aligned}$$

Then, rearranging and sending $\Delta \rightarrow 0$,

$$\begin{aligned} \frac{E(Z_{t+\Delta}^r - Z_t^r | Z_t)}{\Delta} &= Z_t \sum_{k=-1}^n (Z_t + k)^r b_k + o(\Delta) \\ \dot{M}_r | Z_t &= Z_t \sum_{k=-1}^n (Z_t + k)^r b_k \\ &= \sum_{k=-1}^n b_k \sum_{j=0}^r \binom{r}{j} Z_t^{r-j+1} k^j. \end{aligned}$$

The entire $k = 0$ term, which is just $b_0 Z_t^{r+1}$, cancels all of the $j = 0$ terms from the other choices of k , so that we're left with

$$\dot{M}_r | Z_t = \sum_{k=-1, 1, \dots, n} b_k \sum_{j=1}^r \binom{r}{j} Z_t^{r-j+1} k^j.$$

Taking expectations of both sides,

$$\begin{aligned} \dot{M}_r &= \sum_{k=-1, 1, \dots, n} b_k \sum_{j=1}^r \binom{r}{j} M_{r-j+1} k^j \\ &= \sum_{j=1}^r \binom{r}{j} M_{r-j+1} \sum_{k=-1, 1, \dots, n} b_k k^j \\ &= \sum_{j=1}^r \binom{r}{j} M_{r-j+1} E(X_t^j), \end{aligned}$$

as desired. □

Lemma 4. *Hypotheses (H10) and (H11) imply that $\Gamma(t)$ is asymptotically equal to*

$$\int_0^t \frac{g_s''(1)}{2M(s)} ds.$$

Proof. Using Lemma 3, we calculate

$$\begin{aligned} \int_0^t \frac{g_s''(1)}{2M(s)} &= \int_0^t \frac{E(X_s^2) + E(X_s)}{2M(s)} ds \\ &= \Gamma(t) - \int_0^t \frac{-E(X_s)}{2M(s)} ds \\ &= \Gamma(t) - \int_0^t \frac{d}{ds} \frac{1}{2M(s)} ds \\ &= \Gamma(t) - \frac{1}{2M(t)} + \frac{1}{2}. \end{aligned}$$

Hypotheses (H10) and (H11) tell us that this last expression is dominated by the $\Gamma(t)$ term at large times, as desired. \square

Lemma 5. *Hypotheses (H7) and (H9) imply that there exist $T, \eta > 0$ such that for all $s \geq T$ and $t \geq s + 1$, we have $1 - \phi(s, t) \geq \eta$.*

Lemma 5 says that once we reach a point where hypothesis (H7) kicks in, we expect to have some fixed minimal probability of extinction after a fixed time increment.

Proof. For any fixed s , we know that $\phi(s, t)$ is increasing in t . Thus it suffices to prove the claim for $t = s + 1$. By hypothesis (H7), pick T large enough that

$$C \equiv \inf_{x \geq T} b_{-1}(x)$$

is nonzero, and let $D \equiv \sup_{x \geq T} b_1(x) + \dots + b_n(t)$, which by hypothesis (H9) is finite. Define a new birth-and-death process \tilde{Z}_t with constant death rate $\tilde{b}_{-1} = C$ and birth rates $\tilde{b}_1 = \dots = \tilde{b}_n = D$. Then for all times greater than T , \tilde{Z}_t has a smaller death rate and higher birth rates than Z_t , and so the corresponding probability of survival $\tilde{\phi}(s, t)$ is smaller than $\phi(s, t)$. Since the rates of \tilde{Z}_t are constant, we know that $\tilde{\phi}(s, t) = \tilde{\phi}(t - s)$ depends only on the time elapsed, and so picking $\eta = 1 - \tilde{\phi}(1)$ gives our desired uniform bound. \square

Lemma 6. *Hypotheses (H9) and (H11) implies that $\Gamma(t) - \Gamma(t - 1)$ is $o(\Gamma(t))$.*

Lemma 6 will be used to take care of some dangling terms from requiring $t \geq s + 1$ in Lemma 5.

Proof. A first order approximation of Γ gives us

$$\begin{aligned} \Gamma(t) &= \Gamma(t - 1) + \frac{E(X_\xi^2)}{2M(\xi)}, \quad \xi \in (t - 1, t) \\ &\leq \Gamma(t - 1) + \frac{C}{M(\xi)}. \end{aligned}$$

Hypothesis (H11) implies that $\frac{C}{M(\xi)}$ is $o(\Gamma(t))$, and the result follows. \square

3.5 Proof of Theorem 7

As in Theorem 2, Theorem 7 will rely on (4), the preceding lemmas, and a rough lower bound on $g_s''(1 - \phi(s, t))$.

Proof. 1. Using Lemma 4 and hypothesis (H11) with (4) gives us

$$1 \leq \liminf_{t \geq 0} \phi(t) \Gamma(t).$$

2. We expand $g_s''(1 - \phi(s, t))$ by its definition:

$$g_s''(1 - \phi(s, t)) = 2b_1(s) + 6b_2(s)(1 - \phi(s, t)) + \dots + n(n+1)b_n(s)(1 - \phi(s, t))^{n-1}.$$

Consequently, Lemma 5 and hypothesis (H8) imply that there exists a T large enough that

$$\inf_{s \geq T} g_s''(1 - \phi(s, t)) > 0$$

for $t \geq s + 1$. Hypothesis (H9) also implies

$$\sup_{s \geq 0} g_s''(1 - \phi(s, t)) < \infty.$$

Using hypothesis (H10), Lemma 6, hypothesis (H9), and the lower bound on $g_s''(1 - \phi(s, t))$ in that order, we find for $t > T$,

$$\begin{aligned} \Gamma(t) &\leq C \int_T^t \frac{E(X_s^2)}{2M(s)} ds \\ &\leq C \int_T^{t-1} \frac{E(X_s^2)}{2M(s)} ds \\ &\leq C \int_T^{t-1} \frac{g_s''(1 - \phi(s, t))}{2M(s)} ds \\ &\leq C \int_0^t \frac{g_s''(1 - \phi(s, t))}{2M(s)} ds. \end{aligned}$$

Inserting this into (4) and using hypothesis (H11), we find

$$\limsup_{t \geq 0} \phi(t) \Gamma(t) \leq \limsup_{t \geq 0} \frac{C}{\frac{1}{M(t)\Gamma(t)} + 1} = C.$$

Finally, we combine this inequality with the one from 1. to find

$$1 \leq \liminf_{t \geq 0} \phi(t) \Gamma(t) \leq \limsup_{t \geq 0} \phi(t) \Gamma(t) \leq C.$$

Hypothesis (H10) then implies that $\phi(t) \rightarrow 0$.

□

3.6 Proof of Theorem 8

As with Theorem 3, Theorem 8 relies on a more careful estimate of $g_s''(1 - \phi(s, t))$ by splitting the sum in (4) into regions where $\phi(s, t)$ is close to 0 and far away from 0. Once again, we will show that the region where $\phi(s, t)$ is far away from 0 contributes negligibly to the integral.

Proof. 1. Applying the inequality $\phi(t) \leq \frac{C}{\Gamma(t)}$ to the process started at time s gives us

$$\phi(s, t) \leq C \frac{1}{\int_s^t \frac{E(X_u^2)}{2M(u)/M(s)} du} = C \frac{1}{M(s)(\Gamma(t) - \Gamma(s))}.$$

Here we used the fact that $M(u)/M(s)$ is the mean of the process starting at time s .

2. By hypothesis (H9), pick L such that $g_t'''(1) \leq L$ for all t . By hypothesis (H8), pick $T, \eta > 0$ such that $\inf_{t \geq T} g_t''(1) > \eta$. Let $\varepsilon > 0$ and define

$$S(t) = \min\{s : \phi(s, t) \geq \varepsilon\eta/L\}.$$

$S(t)$ is the time that divides our regions where $\phi(s, t)$ is small (when $s < S(t)$), and where $\phi(s, t)$ is large (when $s \geq S(t)$). First note that $S(t) \rightarrow \infty$ as $t \rightarrow \infty$, because extinction is sure. In the case $T < s < S(t)$, the mean value theorem tells us

$$\begin{aligned} g_s''(1 - \phi(s, t)) &= g_s''(1) - g_s'''(\xi)\phi(s, t), \quad \xi \in (1, 1 - \phi(s, t)) \\ &\geq g_s''(1) - L\phi(s, t) \\ &= g_s''(1)(1 - L\phi(s, t)/g_s''(1)) \\ &\geq g_s''(1)(1 - L\phi(s, t)/\eta) \\ &\geq (1 - \varepsilon)g_s''(1). \end{aligned}$$

3. At large t , (4) and our result in 2. tell us

$$\begin{aligned} \frac{1}{\phi(t)\Gamma(t)} &\geq \frac{1}{\Gamma(t)} \int_0^t \frac{g_s''(1 - \phi(s, t))}{2M(s)} ds \\ &\geq \frac{1}{\Gamma(t)} \int_T^{S(t)} \frac{g_s''(1 - \phi(s, t))}{2M(s)} ds \\ &\geq \frac{1}{\Gamma(t)} (1 - \varepsilon) \int_T^{S(t)} \frac{g_s''(1)}{2M(s)} ds \end{aligned}$$

Hypothesis (H10) lets us ignore the loss of the $[0, T]$ interval, so Lemma 4 tells us that

$$\begin{aligned} \frac{1}{\Gamma(t)} (1 - \varepsilon) \int_T^{S(t)} \frac{g_s''(1)}{2M(s)} ds &= \frac{1}{\Gamma(t)} (1 - \varepsilon) \Gamma(S(t)) + o\left(\frac{\Gamma(S(t))}{\Gamma(t)}\right) \\ &= \frac{\Gamma(S(t))}{\Gamma(t)} - \varepsilon \frac{\Gamma(S(t))}{\Gamma(t)} + o(1) \\ &\geq \frac{\Gamma(S(t))}{\Gamma(t)} - \varepsilon + o(1) \\ &= (1 - \varepsilon) - \frac{\Gamma(t) - \Gamma(S(t))}{\Gamma(t)} + o(1). \end{aligned}$$

We want to show that the Γ term tends to zero.

4. Our result in 1. and the definition of $S(t)$ tell us

$$\begin{aligned}\frac{\Gamma(t) - \Gamma(S(t))}{\Gamma(t)} &\leq \frac{C}{M(S(t))\phi(S(t), t)\Gamma(t)} \\ &\leq \frac{CL}{\varepsilon\eta} \frac{1}{M(S(t))\Gamma(t)} \\ &\leq \frac{CL}{\varepsilon\eta} \frac{1}{M(S(t))\Gamma(S(t))},\end{aligned}$$

which tends to zero because $S(t) \rightarrow \infty$. Combining this with 3., we find

$$\liminf_{t \geq 0} \frac{1}{\phi(t)\Gamma(t)} \geq 1 - \varepsilon,$$

which, along with Theorem 7 part 3., tells us

$$1 \leq \liminf_{t \geq 0} \phi(t)\Gamma(t) \leq \limsup_{t \geq 0} \phi(t)\Gamma(t) \leq \frac{1}{1 - \varepsilon},$$

for every $\varepsilon > 0$, completing the proof. □

3.7 Proof of Theorem 9

In Theorem 9 we depart more significantly from the discrete setting. We still prove the claim by induction, but now we rely on the differential formula of Lemma 3.

Proof. 1. We proceed by induction. The base case $r = 1$ is clear. Now assume that

$$\frac{M_j(t)}{M^j(t)} = j!\Gamma(t)^{j-1} + o(\Gamma(t)^{j-1}) \quad (j = 1, \dots, r-1).$$

Recall from Lemma 1 that $\dot{M} = ME(X_t)$. Using Lemma 1, we find

$$\begin{aligned}(M_r/M^r)'(t) &= \frac{1}{M^r} \sum_{j=1}^r \binom{r}{j} M_{r-j+1} E(X_t^j) - r \frac{M_r}{M^{r+1}} \dot{M} \\ &= \frac{1}{M^r} \sum_{j=1}^r \binom{r}{j} M_{r-j+1} E(X_t^j) - \frac{1}{M^r} r M_r E(X_t) \\ &= \frac{1}{M^r} \sum_{j=2}^r \binom{r}{j} M_{r-j+1} E(X_t^j).\end{aligned}$$

Then our inductive hypothesis tells us

$$(M_r/M^r)'(t) = \sum_{j=2}^r \binom{r}{j} E(X_t^j) (r-j+1)! \frac{\Gamma(t)^{r-j}}{M(t)^{j-1}} + o\left(\frac{\Gamma(t)^{r-j}}{M(t)^{j-1}}\right).$$

Hypothesis (H11) tells us that the term according to $j = 2$ dominates the sum at large times, and thus

$$\begin{aligned}(M_r/M^r)'(t) &= \frac{r(r-1)}{2} E(X_t^2) \frac{(r-1)! \Gamma(t)^{r-2}}{M(t)} + o\left(\frac{\Gamma(t)^{r-2}}{M(t)}\right) \\ &= r!(r-1) \Gamma(t)^{r-2} \frac{E(X_t^2)}{2M(t)}.\end{aligned}$$

2. On the other hand, differentiating $\Gamma(t)^{r-1}$ yields

$$\begin{aligned}(\Gamma^{r-1})'(t) &= (r-1) \Gamma(t)^{r-2} \dot{\Gamma}(t) \\ &= (r-1) \Gamma(t)^{r-2} \frac{E(X_t^2)}{2M(t)}.\end{aligned}$$

Combining this with 1., we find

$$(M_r/M^r)'(t) = r!(\Gamma^{r-1})'(t) + o((\Gamma^{r-1})'(t)),$$

and integrating both sides yields the desired result. □

3.8 Proof of Theorem 6

As in the discrete setting, proving the main theorem is now just a matter of putting together the pieces of the supporting theorems.

Proof. The moments of the conditioned, scaled process can be written as

$$\begin{aligned}E\left(\left(\frac{Z_t}{E(Z_t|Z_t \neq 0)}\right)^r \middle| Z_t \neq 0\right) &= \frac{E(Z_t^r | Z_t \neq 0)}{E(Z_t | Z_t \neq 0)^r} \\ &= \frac{M_r(t)/\phi(t)}{(M(t)/\phi(t))^r} \\ &= \frac{M_r(t)\phi(t)^{r-1}}{M(t)^r}.\end{aligned}$$

Theorems 8 and 9 tell us

$$\lim_{t \rightarrow \infty} \frac{M_r(t)\phi(t)^{r-1}}{M(t)^r} = \lim_{t \rightarrow \infty} \frac{M_r(t)}{M(t)^r \Gamma(t)^{r-1}} = r!.$$

Thus the r -th moment of the conditioned, scaled process tends to $r!$ and so the process converges to an exponential with parameter 1 in moments. □

3.9 Proof of Theorem 10

Proof. First we calculate

$$\int_0^t \frac{E(X_s)}{M(s)} ds = - \int_0^t \frac{d}{ds} \frac{1}{M(s)} ds = 1 - \frac{1}{M(t)}.$$

Thus, $M(t) = \left(1 - \int_0^t \frac{E(X_s)}{M(s)} ds\right)^{-1}$. Let $\varepsilon > 0$ and pick T large enough that $|E(X_t)| < \varepsilon$ for all $t \geq T$. Then using hypotheses (H8) and (H9) to bound $E(X_t^2)$ from below and above, when $t \geq T$ we calculate

$$\begin{aligned} \frac{1}{\Gamma(t)M(t)} &\leq \frac{1}{\Gamma(t)} + C \frac{\int_0^t \frac{|E(X_s)|}{M(s)} ds}{\int_0^t \frac{1}{M(s)} ds} \\ &\leq \frac{1}{\Gamma(t)} + \frac{C \int_0^T \frac{E(X_s)}{M(s)} ds}{\int_0^t \frac{1}{M(s)} ds} + \frac{C \int_T^t \frac{E(X_s)}{M(s)} ds}{\int_T^t \frac{1}{M(s)} ds} \\ &\leq \frac{1}{\Gamma(t)} + \frac{C}{\Gamma(t)} + C\varepsilon. \end{aligned}$$

The first two terms tend to zero by hypothesis (H10), and the result follows. \square

References

- [1] K. B. Athreya and P. E. Ney, Branching Processes, 1-24. Springer, New York (1972)
- [2] A. Agresti, On the Extinction Times of Varying and Random Environment Branching Processes, J. Appl. Prob., 12, 39-46 (1975)
- [3] V. P. Cistjakov and N. P. Markova, On Some Theorems for Inhomogeneous Branching Processes, Dokl. Akad. Nauk., 147, 317-320 (1962)
- [4] P. Jagers, Galton-Watson Processes in Varying Environments, J. Appl. Prob., 11, 174-178 (1974)
- [5] M. Jirina, Extinction of Non-Homogeneous Galton-Watson Processes, J. Appl. Prob., 13, 132-137 (1976)
- [6] D. G. Kendall, On the Generalized “Birth-and-Death” Process, The Annals of Mathematical Statistics, 19, 1-15 (1948)